UNIVERSITI TEKNOLOGI MARA

COPY NUMBER VARIATIONS OF 'ORANG ASLI' (NEGRITO) FROM PENINSULAR MALAYSIA

SITI SHUHADA MOKHTAR

Thesis submitted in fulfillment
of the requirements for the degree of
Master of Science

Faculty of Medicine

April 2015

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	: Siti Shuhada Binti Mokhtar
Student I.D. No.	: 2010692186
Programme	: Master of Science (MD 780)
Faculty	: Medicine
Thesis Tittle	: Copy Number Variations of 'Orang Asli'(Negrito) From Peninsular Malaysia
Signature of Student	: .....
Date	: April 2015

ABSTRACT

Copy number variation (CNV) has been acknowledged as a major contributor to the human genome diversity. This variability covers approximately about 15% of the entire human genome. To date, it is reported that CNV plays an important role in the association of common and complex diseases and also in determining the phenotypes. Nevertheless, CNV data from diverse populations is still limited and not saturated especially population from South East Asia. This study represents the first investigation of CNV in the indigenous populations from Peninsular Malaysia, of particular interest, the Negrito. Ninety-seven Negrito samples were recruited, of which 50 unrelated Negrito samples were genotyped using the Affymetrix SNP 6.0 microarray. CNVs were then called by two independent algorithms namely, Genotyping Console and Nexus CNV. Subsequently the CNVs being called by both algorithms were considered stringent hence included to the subsequent analyses. A total of 643 stringent CNVs were identified, comprising 217 gains and 426 losses. These stringent CNVs were then matched with the publicly available datasets including DGV, HapMap3 and SGVP, and identified 57 putative novel and Negrito specific CNVs, consisting of 28 CNVRs. These included 15 gains and 13 losses. Analysis of gene ontology revealed that genes within these CNVs were enriched in the immune system ($p < 0.001$). In view of the small population size, relative isolation and semi-nomadic lifestyles of this community, it is speculated that these CNVs may be attributed to recent local adaptation of Negritos from Peninsular Malaysia. Nonetheless other factors such as genetic drift should not be ruled out. This study offers a preliminary effort at an extent to which rare variants shape risk of common disease. More saturated CNV map is needed and should be carried out through large-scale next generation sequencing.

TABLE OF CONTENTS